

**APPLICATION  
FOR  
UNITED STATES LETTERS PATENT**

**APPLICANT NAME:** Nowicki et al

**TITLE:** EFFICIENT SCALING IN TRANSFORM DOMAIN

**DOCKET NO.** BLD920030011US1 (IRA-10-5737)

**INTERNATIONAL BUSINESS MACHINES CORPORATION**

CERTIFICATE OF MAILING UNDER 37 CFR 1.10

I hereby certify that, on the date shown below, this correspondence is being deposited with the United States Postal Service in an envelope addressed to the Mail Stop Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 as "Express Mail Post Office to Addressee"

Mailing No. EL 897036135US  
on SEPTEMBER 30, 2003

Kathleen Kent  
Name of person mailing paper

Signature: Kathleen Kent

Date: 9/30/2003

# **EFFICIENT SCALING IN TRANSFORM DOMAIN**

## **CROSS-REFERENCE TO RELATED APPLICATIONS**

The present application is related to the following co-pending and commonly assigned United States patent applications, which are hereby incorporated by reference in their respective entirety: serial number \_\_\_\_\_ filed on \_\_\_\_\_ by Mitchell et al. for “System and Method for Enabling Multiple Signed Independent Data Elements per Register” (IBM docket number BLD920000060); serial number 09/570,382 filed on May 12, 2000 by T. J. Trenary et al. for “Method and Apparatus for the Scaling Up of Data”; serial number 09/570,849 filed on May 12, 2000 by J. L. Mitchell et al. for “Method and Apparatus for the Scaling Down of Data”; serial number \_\_\_\_\_ filed on \_\_\_\_\_ by Trelewicz et al. for “Faster Transforms Using Scaled Terms” (IBM docket number BLD920000059); serial number \_\_\_\_\_ filed on \_\_\_\_\_ by Trenary et al. for “Reduction of N DCT blocks into One Block” (IBM docket number BLD919990036); and serial number \_\_\_\_\_ filed on \_\_\_\_\_ by Tomasz Nowicki et al for “Method and System for Scaling a Signal Sample Rate” (IBM docket number YOR920020113US1).

## **FIELD OF THE INVENTION**

This invention relates in general to data processing, and more particularly to data transforms that use scaled terms. More particularly, the present invention addresses high-end color printer performance for scaling operations.

## **BACKGROUND OF THE INVENTION**

Transforms, which take data from one domain (e.g., sampled data) to another (e.g., frequency space), are used in many signal and/or image processing applications. Such transforms are used for a variety of applications, including, but not limited to data analysis, feature identification and/or extraction, signal correlation, data compression, or data embedding. Many of these transforms require efficient implementation for real time and/or fast execution whether or not compression is used as part of the data processing.

Data compression is desirable in many data handling processes, where too much data is present for practical applications using the data. Commonly, compression is used in communication links, to reduce transmission time or required bandwidth. Similarly, compression is preferred in image storage systems, including digital printers and copiers, where "pages" of a document to be printed may be stored temporarily in memory. Here the amount of media space on which the image data is stored can be substantially reduced with compression. Generally speaking, scanned images, i.e., electronic representations of hard copy documents, are often large, and thus make desirable candidates for compression.

In data processing, data is typically represented as a sampled discrete function. The discrete representation is either made deterministically or statistically. In a deterministic representation, the point properties of the data are considered, whereas, in a statistical representation, the average properties of the data are specified. In particular examples referred to herein, the terms images and image processing will be used. However, those skilled in the art will recognize that the present invention is not meant to be limited to processing still images but is applicable to processing different data, such as audio data, scientific data, video data, sensor data, etc.

In a digital image processing system, digital image signals are formed by first dividing a two-dimensional image into a grid. Each picture element, or pixel, in the grid has associated therewith a number of visual characteristics, such as brightness and color. These characteristics are converted into numeric form. The digital image signal is then formed by assembling the numbers associated with each pixel in the image into a sequence which can be interpreted by a receiver of the digital image signal.

Signal and image processing frequently require converting input data into transform coefficients for the purposes of analysis. Often only a quantized version of the transform coefficients is needed, such as, for example, JPEG/MPEG data compression or audio/voice compression. Many such applications need to be done fast in real time such as the generation of JPEG data for high speed printers.

One compression technology defined in the JPEG standard, as well as other emerging compression standards, is discrete cosine transform (DCT) coding, wherein an input image is divided into many uniform image blocks with data samples in each, typically in an 8x8 array of data samples, to achieve image compression. Images compressed using DCT coding are decompressed using an inverse transform known as the inverse DCT (IDCT). A two-dimensional forward discrete cosine transform (FDCT) function is applied to each block to transform the data samples into a set of transform coefficients to remove the spatial redundancy.

In general, the forward transform will produce real-valued data, not necessarily integers. To achieve data compression, the transform coefficients are converted to integers by the process of quantization. The resulting integers are then passed on for possible further encoding or compression before being stored or transmitted.

The two basic components of an image compression/decompression system are the encoder and the decoder. The encoder compresses the “source” image (the original digital image) and provides an output of compressed data (or coded data). The compressed data may be either stored or transmitted, but at some point are fed to the decoder. The decoder recreates or “reconstructs” an image from the compressed data.

In general, a data compression encoding system may include three basic parts: an encoder model, an encoder statistical model, and an entropy encoder. The encoder model generates a sequence of “descriptors” that is an abstract representation of the image. The statistical model converts these descriptors into symbols and passes them on to the entropy encoder. The entropy encoder, in turn, compresses the symbols to form the compressed data. The encoder may require external tables. That is, tables specified externally when the encoder is invoked. Generally, there are two classes of tables; model tables that are needed in the procedures that generate the descriptors, and entropy-coding tables that are needed by the JPEG entropy-coding procedures.

JPEG uses two techniques for entropy encoding: Huffman coding and arithmetic coding. Similarly to the encoder, the decoder may include basic parts that have an inverse function relative to the parts of the encoder.

JPEG compressed data contains two classes of segments: entropy-coded segments and marker segments. Other parameters that are needed by many applications are not part of the JPEG compressed data format. Such parameters may be needed as application-specific “wrappers” surrounding the JPEG data; e.g., image aspect-ratio, pixel shape, orientation of image, etc.

Within the JPEG compressed data, the entropy-coded segments contain the entropy-coded data, whereas the marker segments contain header information, tables, and other information required to interpret and decode the compressed image data. Marker segments always begin with a “marker”, a unique 2-byte code that identifies the functions of the segment.

To perform a display (or print or audio) operation, it may be necessary for the display device to scale an image to a larger or smaller size. The scaling of the images may be performed as a linear operation. The array of coefficients describing the intensity of the colors of the pixels of the image is transformed to an array of coefficients of the scaled image by a matrix operation.

This transformation may be performed in any representation of the image, but may depend on such a representation. As long as the representation is linear with respect to the pixel values the transformation stays linear.

The scale factor is a number which expresses the ratio of the number of samples in the image before and after the scaling. Usually the scaling is performed block-wise, where the size of the block (which may be the entire signal) is determined by the scale factor, the demanded efficiency of the operation and the quality of the resulting signal. Choosing larger blocks may yield better quality but lesser efficiency, because the larger blocks allows the scaling factor to be approximated more accurately. The scale factors with small integers as denominators and numerators allow smaller blocks, larger integers may force larger blocks.

The emphasis of the present invention is addressing high-end color printer performance for scaling operations. Currently, scaling a continuous tone JPEG image has a strong undesirable effect on the throughput of the printer. Conventional prior art image

reduction processes typically involve doing an IDCT transform on each 8x8 DCT block to create real domain data (64 samples), reducing the image in the pixel domain, and then doing a FDCT to return to the DCT domain. The main problem with this approach is that it is computationally expensive. For full-page images the IDCT and FDCT calculations alone could exceed the total processing time available, particularly if the images are being reduced down to make them fit on a page.

In one reference incorporated above, "Reduction of N DCT blocks into One Block" by Trenary et al., a solution has been developed wherein one-dimensional DCT domain reduction methods merge N blocks along one dimension into one block, resulting in a significant transactional savings. This approach offers computationally efficient advantages in 1/n "downscaling" operations. However, where the same method and system is utilized in "up-scaling" operations, extra computational cycles are required, reducing efficiency advantages. Moreover, the extra computational cycles introduce additional opportunities for the occurrence of errors through additional "round-off" steps.

One area where both data transform and scaling operations are required is high impression-per-minute ("ipm") printing during "contone" (continuous tone; e.g., photographic) image scaling. The criticality of the problem increases as the printer speed is increased. What is needed is a computationally efficient system and method to provide transform and scaling operations in data processing, and more particularly in data transform operations that use scaled terms. More particularly, an improved system and method is required to address high-end color printer performance for scaling operations.

### **SUMMARY OF THE INVENTION**

A method and system for efficient scaling in the transform domain, wherein transform coefficient data is provided as an input to a data processing system and scaled in the transform domain by application of a combined matrix. Some embodiments utilize discrete cosine transform data. One embodiment of the invention generates a combined matrix for one-dimensional scaling by selecting a rational scaling factor and matrix dimension value, generating a matrix with some zero values, applying a one-dimensional inverse transform, regrouping, and applying a one-dimensional forward transform. One

application of the invention performs up-scaling operations, and another performs down-scaling operations. The invention also provides for two-dimensional scaling by selecting horizontal and vertical scaling parameters and generating first and second combined matrices responsive to the parameters and combining them into a single combined matrix.  
5 The invention may also incorporate a predetermined cost function.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 is a diagram showing a block structure of scaling matrices according to the present invention with a down-scaling factor of  $1/2$ .

Figure 1A is a diagram showing the content of the blocks of the scaling matrices  
10 of Figure 1.

Figure 2 is a diagram showing a block structure of scaling matrices according to the present invention with a down-scaling factor of  $1/4$ .

Figure 2A is a diagram showing the content of the blocks of the scaling matrices  
15 of Figure 2.

Figure 3 is a diagram showing a block structure of scaling matrices according to the present invention with an up-scaling factor of 2.

Figure 3A is a diagram showing the content of the blocks of the scaling matrices  
20 of Figure 3.

Figure 4 is a diagram showing a block structure of scaling matrices according to the present invention with an up-scaling factor of 4.

Figure 4A is a diagram showing the content of the blocks of the scaling matrices  
25 of Figure 4.

Figure 5 illustrates an article of manufacture comprising a computer usable medium having a computer readable program according to the present invention embodied in said medium, as implemented in a printer.

### **DESCRIPTION OF THE PREFERRED EMBODIMENTS**

A method and system for efficient scaling in the transform domain when transform domain data is provided as an input to a system, comprising scaling the transform domain data input in one combined matrix operation step in the transform

domain. The invention relates in general to data processing, and more particularly to data transforms that use scaled terms. In illustrative embodiments of the present invention described herein the intended application is for high-end color printer performance for scaling operations, specifically a system and method that speeds scaling of JPEG images by using the structure of the scaling matrices, combined with the structure of the FDCT and IDCT transforms employed by JPEG, to create one composite transform that performs the scaling and "repackaging" of DCT coefficients into 8x8 blocks.

It is to be understood that although the present embodiments are intended for JPEG image applications that the present invention is not limited to such applications. It will be readily apparent to one skilled in the art that the present invention can be readily adapted to a wide variety of data processing tasks that require efficient scaling in the transform domain when the transform domain data is provided as an input to a system.

Because the contone images are received at the printer in JPEG format, they are already in the DCT domain, making this method very efficient, since it eliminates the need to transform the data back to the pixel domain prior to manipulation. Entropy coding must be removed from the data prior to application of an algorithm by the present invention; however, entropy coding must necessarily be removed from the data before subsequent processing in the printer anyway, so this requirement does not introduce additional operations.

Other prior art references, such as "Method and Apparatus for the Scaling Up of Data" by Trenary et al. and "Method and Apparatus for the Scaling Down of Data" by Mitchell et al. (both previously incorporated by reference) teach "scaling up" and "scaling down" through matrix operations. However, both of these references teach systems and methods wherein the actual cosines must be kept with the transform constants. What is important in the present invention is a computation-efficient implementation of the constants in the scaling matrix without the actual cosines.

An important advantage of the present invention is in how the matrices are constructed. Matrices used in prior art scaling use floating point or simple fixed point approaches, while the present invention uses the integer methods to directly address



computational complexity. The integer methods utilized are taught in “Faster Transforms Using Scaled Terms” by Trelewicz et al, previously incorporated by reference. As taught by the present invention, integer computational optimization can also be used to reduce cache misses on computer system devices, such as modems and pipelined processors; to make efficient field programmable gate array (FPGA) hardware implementations for hardware systems; and to reduce computational cycles on a range of embedded processors for pervasive applications. Furthermore, contrast and image quality feed directly into the cost functions used for optimization of the matrices for computation, and are flexible for a range of applications.

The present invention may be described as an implementation of “one scaling transform”, which can perform inverse transforms, scaling, and forward transforms combined into one matrix operation on multiple transform coefficient blocks. Thus scaling examples according to the present invention become specific cases of combined linear operations.

The present invention provides for significant advantages in both down-scaling and up-scaling of contone images.

Down-scaling. Scaling an image down requires low-pass-filtering of the image to avoid "aliasing", an effect in sampled signals and images where high frequency content becomes low frequency noise when the high frequency components exceed the Nyquist frequency of the resampled signal. The “Nyquist limit” is commonly defined as the highest frequency of input signal that can be correctly sampled, equal to half of the sampling frequency. However, in the DCT domain, the deletion of high-frequency coefficients (replacement with zero) is equivalent to high-quality low-pass-filtering. At this point, the zero high-frequency coefficients can be removed from the DCT block, forming a smaller block, for example “ $(n) \times (n)$ ”. When an  $(n) \times (n)$  IDCT is applied to this block, the down-sampled image results. However, it should be noted that such an operation can produce pixel-domain results out-of-range: for example, if the original samples were in the range 0--255, the scaled pixels, after application of the IDCT, can be smaller than 0 and/or larger than 255, requiring some type of operation to bring them

back into range. This effect results from the mathematics of the DCT, and is predictable and reproducible.

Basic matrix structures for down-scaling by  $n/m$  where  $n < m$ ,  $k = g(n)/m$  and  $M = m$ , are as follows:

- 5 (1) The matrix  $P$  is of the form:  $[[ (n)x(n) ] [ (n)x(m-n) ] ]$ , where the  $[(n)x(m-n)]$  matrix is identically zero, but it will act on a matrix of the form  $[[ (m)x(M) ]_1 \dots [ (m)x(M) ]_g ]$ ; so we can assume its form to be:  
 $[[ [ (n)x(n) ] [ (n)x(m-n) ] ]_1 \dots [ [ (n)x(n) ] [ (n)x(m-n) ] ]_g ]$ ,  
10 where each  $[(n)x(m-n)]$  matrix is zero;
- (2) The inverse transform matrix  $d(n)$  of the form  $d(n) = [(n)x(n)]$  acts on the result of (1) (leaving its structure untouched) and then this result is regrouped using the relationship  $k(m) = g(n)$  to produce:  
 $[[ (m)x(M) ]_1 \dots [ (m)x(M) ]_k ]$ ;
- 15 (3) Then the forward transform matrix  $D(m)$  of the form  $D(m) = [(m)x(m)]$  acts on the result of (2).

Under the present invention this process can also be achieved equivalently one dimension at a time: From an initial  $(m)x(m)$  block creating an  $(n)x(m)$  block, repackaging, and then creating an  $(n)x(n)$  block, where  $m$  is the dimension value of an  $(m)x(m)$  matrix. Note that the present invention is illustrated scaling both dimensions  
20 equally. However, since each dimension is done independently, the result could be an  $(n')x(n)$  block where  $n'$  is not equal to  $n$ . Note that the collection of conceptual  $(n) \times (n)$  blocks may be repackaged into a smaller number of  $(m) \times (m)$  blocks as part of the combined matrix operation. One dimension could be scaled up and the other axis scaled  
25 down. As JPEG processing is particularly suited for manipulation of data in  $8 \times 8$  blocks, it is intended that  $m = 8$  for JPEG imaging applications. However, other values of  $m$  may be selected for use with the present invention.

For example, an  $n/m$  scaling down along one axis may be performed according to the present invention through the following steps:

- (a) Select  $g$  as the smallest integer such that  $(ng)/m$  is an integer  $k$ ;
- (b) Define  $X$  to be an  $(mg) \times (m)$  matrix of DCT coefficients formed by taking  $g$   $(m) \times (m)$  blocks;
- (c) Define  $p$  as an  $(ng) \times (m)$  matrix, built of  $g$  blocks of  $(n) \times (m)$  equal matrices, which when applied will reject the highest  $m-n$  frequencies of each  $(m) \times (m)$  block along the axis being scaled down one;
- (d) Define  $d_g(n)$  as an  $(ng) \times (n)$  IDCT transform (which is implemented by a matrix), consisting of  $g$  blocks of  $(n) \times (n)$  IDCT transforms; and
- (e) Define  $D_g(n)$  as an  $(ng) \times (m)$  FDCT matrix, consisting of  $ng/m$  blocks of  $(m) \times (m)$  FDCT transforms for repackaging  $d_g(n)pX$  into  $ng/m$   $(m) \times (m)$  DCT blocks.
- (f) Define  $S = D_g(n)d_g(n)p$ .

The  $SX$  operation outputs  $k$  blocks from the original  $g$  blocks.  $S$ , the combined matrix is a “sparse matrix”: it has many zero entries. In one embodiment of the present invention the algorithm process taught by “Faster Transforms Using Scaled Terms” by Trelewicz et al, previously incorporated by reference, is employed to find optimal integer representations for the  $S$  matrix constants, with the common denominator  $q$  for the integer approximation operations being adjusted so that the resulting contrast is within a predetermined range of the original 100% contrast.

The common denominator  $q$  may be found according to the methods taught by J. Q. Trelewicz, Michael T. Brady and Joan L. Mitchell in “Efficient Integer Implementations For Faster Linear Transforms”, in Proc. of 35th Asilomar Conf. on Signals, Systems, and Computers 2001 , (Pacific Grove, CA), 4-7 Nov 2001. There the common denominators used for the subtransforms are chosen according to a cost function tailored to the specific application and implementation architecture. For example, the cost function may take into account the number of bits available in the hardware for calculation, the amount of error that can be tolerated in a calculation, and the resulting complexity of the calculation on that architecture. Thus, the implementation of the transform with smaller constants can reduce the need for memory accesses, thus reducing

cache misses. Although the present embodiment utilizes DCT transform structures, this architecture also works for other transforms. Moreover, even greater flexibility is provided using simultaneous rational approximations (i.e., a common denominator  $q$ ) to all of the constants in a subtransform, since the simultaneous representations can be found in accordance with the cost function. In architectures preferring shifts and additions to multiplications, the numerators of the rational approximations may be viewed as polynomials in powers of 2 with plus/minus 1 or 0 coefficients.

In one embodiment a cost function finds simultaneous representations (numerators) with the smallest number of common power-of-2 terms; i.e., the set of power-of-2 terms in all of the polynomials in the representations of a subtransform is as small as possible. This formulation allows the power-of-2 terms to be grouped, so that the number of operations in the shift-and-add transform can be reduced. Using this cost function adjustment method for the integer approximation operation, the predetermined range may be chosen in the present invention so that representations for the matrix  $S$  cannot produce scaled DCT coefficients outside the preferred range. One preferred predetermined range of 80% to 120% of original contrast produces high-quality results.

Because of the way in which  $S$  is represented per the cost function, it is suited for efficient implementation in software or hardware, using the parallel processing methods of Mitchell et al., "System and Method for Enabling Multiple Signed Independent Data Elements per Register", previously incorporated by reference.

Scaling on the other axis is an extension of this method, by using the transpose of the matrices. Figure 1 illustrates an example of down-scaling by  $1/2$  according to the present invention, and Figure 2 illustrates an example of down-scaling by  $1/4$  according to the present invention. Both examples are more fully discussed below.

Up-scaling. Scaling an image up cannot increase the frequency content of the image; i.e., only the lower frequencies already present in the image can be present in the larger-scale image, since no additional information is present in the image. Thus, in a similar manner to the down-scaling mentioned above, up-scaling can be achieved by increasing the size of the DCT block by inserting zero coefficients at the high frequencies

to create, say, a  $(N) \times (N)$  DCT block. An  $(N) \times (N)$  IDCT then results in the up-scaled image.

Basic matrix structures for up-scaling by  $N/m$  where  $m < N$  and  $k = g(N)/m$ , are as follows:

- 5 (a) The matrix  $P$  is of the form:  $[(m) \times (m)][(N-m) \times (m)]$ , where the  $[(N-m) \times (m)]$  matrix is all zeros, but it will act upon a matrix of the form  $[(m) \times (M)]_1 \dots [(m) \times (M)]_g$ ; so we can assume its form to be:  
 $[[[(m) \times (M)][(N-m) \times (M)]]_1 \dots [[[(m) \times (M)][(N-m) \times (M)]]_g]$ ,  
10 where each  $[(N-m) \times (M)]$  submatrix is identically zero;
- (b) The inverse transform matrix  $d(N)$  of the form  $d(N) = [(N) \times (N)]$  acts on the result of (a) (leaving its structure unchanged) and then this result is regrouped using the relationship  $g(N) = k(m)$  as:  
 $[[[(m) \times (M)]_1 \dots [(m) \times (M)]_k]$ ;
- 15 (c) Then the forward transform matrix  $D(m)$  of the form  $D(m) = [(m) \times (M)]$  acts on the result of (b).

In the same manner as scaling down, this process can also be achieved equivalently one dimension at a time; i.e., from a  $(m) \times (m)$  block creating an  $(N) \times (m)$  block, repackaging, and then creating an  $(N') \times (N)$  block where  $N'$  and  $N$  are not necessarily equal. Note that the collection of conceptual  $(N) \times (N)$  blocks may be repackaged into a larger number  $(m) \times (m)$  blocks as part of the combined matrix operation. Therefore an  $N/m$  scaling up according to the present invention may be performed as follows:

- (a) select  $g$  as the smallest integer wherein  $Ng/m$  is an integer  $k$ ;
- 25 (b) define  $X$ , to be an  $(mg) \times (m)$  matrix of DCT coefficients formed by taking  $g$   $(m) \times (m)$  blocks;
- (c) define  $P$ , an  $(Ng) \times (m)$  matrix, which thereby inserts zeros at the  $N-m$  high frequencies in one dimension in each  $(m) \times (m)$  DCT-block;

- (d) define  $d_g(N)$ , an  $(Ng) \times (N)$  IDCT matrix consisting of  $g$  blocks of  $(N) \times (N)$  IDCT transforms; and
- (e) define  $D_g(n)$ , an  $(Ng) \times (m)$  DCT matrix, consisting of  $Ng/m$  blocks of  $(m) \times (m)$  FDCT transforms for repackaging  $d_g(N)PX$  into  $Ng/m$   $(m) \times (m)$  DCT blocks.
- (f) Define  $S = D_g(N)d_g(N)P$ , the combined matrix.

The  $SX$  operation outputs  $k$  blocks from the original  $g$  blocks.  $S$  is also a sparse matrix. Figure 3 illustrates an example of up-scaling by a factor of 2 according to the present invention, and Figure 4 illustrates an example of up-scaling by a factor of 4 according to the present invention. Both examples are more fully discussed below.

Examples of the present invention. Now with reference to Figures 1 and 1A, the structure of the matrix 2, labeled  $S_{1/2}$ , of scaling by factor 1/2 is illustrated. It has eight rows and sixteen columns and is split into four blocks 4:  $A$ ,  $Z$ ,  $A'$ ,  $Z$ , wherein each block 4 has eight rows and four columns. The second block 4b and the fourth block 4d, labeled  $Z$ , are equal and have all entries zero. The entries of the first block 4a labeled  $A$  are shown in Figure 1A. The third block 4c labeled  $A'$  is a “checkerboard” matrix of the entries of the first block 4a labeled  $A$ . The generation of the checkerboard matrix block 4c is conventional, wherein block 4a is indexed by counting from 1 in both the horizontal and vertical directions. For example, assume a  $2 \times 2$  matrix  $M$  with the following entries:

$$\begin{matrix} W & X \\ Y & Z \end{matrix}$$

$W$  is at 1,1;  $X$  at 1,2;  $Y$  at 2,1; and  $Z$  at 2,2. In order to generate a checkerboard matrix  $M'$ , the corresponding entries of matrix  $M$  are adjusted wherein the sign of every element with an “odd,even” or “even,odd” index is flipped, but the “even,even” or “odd,odd” index are not flipped. So here, we would flip  $X$  and  $Y$ , but not  $W$  or  $Z$ , and accordingly matrix  $M'$  has the following entries:

$$\begin{matrix} W & -X \\ -Y & Z \end{matrix}$$

Now with reference to Figure 2 and 2A, the matrix 10 labeled  $S_{1/4}$  is provided to illustrate down-scaling by factor 1/4 according to the present invention. Matrix 10 has eight rows and thirty-two columns. It is split into 8 blocks 12:  $A$ ,  $Z$ ,  $B$ ,  $Z$ ,  $A'$ ,  $Z$ ,  $B'$ , and  $Z$ . The second, fourth, sixth and eighth blocks 12 labeled  $Z$  are equal, having eight rows and

six columns with all entries zero. The blocks 12 labeled  $A$ ,  $A'$ ,  $B$ ,  $B'$  each have eight rows and two columns. The entries of blocks 12 labeled  $B$  and  $B'$  are the same except for sign changes according to the checkerboard pattern process described above, and the entries of blocks 12 labeled  $A$  and  $A'$  are also the same except for sign changes according to the said checkerboard pattern process.

Figures 3 and 3A illustrate up-scaling by a factor of two according to the present invention. Figure 3A shows an 8x8 matrix 20  $A$ . The matrix 22  $A'$  is generated from matrix 20 through the checkerboard process described above, and the two matrixes 20 and 22 are combined to form composite matrix 24, with sixteen rows and eight columns. The matrix 24 has a block structure, with the first eight rows forming the matrix  $A$  and the last eight rows forming the 8x8 matrix called  $A'$ . The entries of  $A'$  and  $A$  are the same except for the sign change in the checkerboard pattern as described above.

Figures 4 and 4A illustrate up-scaling by a factor of four according to the present invention. Matrix 30 is a 32x8 matrix, with thirty-two rows and eight columns, and has a block structure. The first eight rows are formed by the 8x8 matrix 32 labeled  $A$ . Rows nine through sixteen of matrix 30 are formed by 8x8 matrix 34 labeled  $B$ . Rows seventeen through twenty-four of matrix 30 are formed by the 8x8 matrix 36  $B'$ , which is generated from matrix 34 labeled  $B$  through the checkerboard process described above. And lastly, the last eight rows of matrix 30 are formed by the 8x8 matrix 38 labeled  $A'$ . Again, the entries of matrix 38  $A'$  and matrix 32  $A$  are the same except for the sign change in the checkerboard pattern.

In Figures 1A, 2A, 3A and 4A the entries of the labeled blocks are the fractions with denominator 32, and the multiplication by such matrices is treated as multiplication by the numerators which are integers 1,2,..,31. Each multiplication is implemented as a sequence of shifts (i.e., multiplied by a power of 2), additions, or subtractions according to the sign of the entry and the methods taught in "System and Method for Enabling Multiple Signed Independent Data Elements per Register" by Mitchell et al., previously incorporated by reference. The division by 32 of the resulting sum is implemented as a shift right, after the calculation with the numerator is completed. The checkerboard

symmetry of signs is exploited by precalculation of sums and differences of pairs of input data. For the description of the invention we assume that the scaling is done in the rows and hence the data is represented by 16 (for scale factor 1/2) or 32 (for scale factor 1/4) rows of dequantized data. The same method is employed for scaling the columns of data. Equivalently, columns could be scaled first, and rows second.

Referring now to Figure 5, an embodiment of the invention described above may be tangibly embodied in a computer program residing on a computer-readable medium or carrier 490. The medium 490 may comprise one or more of a fixed and/or removable data storage device such as a floppy disk or a CD-ROM, or it may consist of some other type of data storage or data communications device. The computer program may be loaded into the memory 492 to configure the processor 440 for execution. The computer program comprises instructions which, when read and executed by the processor 440 causes the processor 440 to perform the steps necessary to execute the steps or elements of the present invention.

The foregoing description of the exemplary embodiment of the invention has been presented for the purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise forms disclosed. Many modifications and variations are possible in light of the above teaching. It is intended that the scope of the invention be limited not with this detailed description, but rather by the claims appended hereto.